

ANÁLISE DE EFICIÊNCIA DE COMPRESSÃO DE *LIGHT FIELDS* UTILIZANDO O SOFTWARE DE REFERÊNCIA JPEG MULE

**EDUARDO AMARO; IAGO STORCH; DOUGLAS CORRÉA; BRUNO ZATT;
LUCIANO AGOSTINI**

*Universidade Federal de Pelotas
Video Technology Research Group (ViTech)
{eadrosa, icstorch, dscorrea, zatt, agostini}@inf.ufpel.edu.br*

1. INTRODUÇÃO

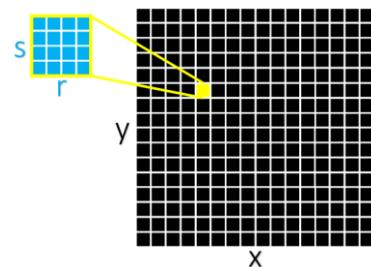
Vídeos digitais estão se tornando cada vez mais populares. Isso se deve ao fato de que dispositivos móveis, como *smartphones* e *tablets*, que são capazes de capturar, reproduzir e disseminar este tipo de mídia, estão se difundindo à escala global. Agregado a isso, serviços de streaming online, como *Twitch*, *Netflix* e *YouTube* estão fazendo parte do dia a dia de grande parte da população, contribuindo para o consumo de vídeos digitais e no aumento do mesmo no tráfego da internet. Estima-se que no ano de 2022, 82% de todo o tráfego de dados na internet seja exclusivamente de vídeos digitais (CISCO, 2019).

Neste cenário, cada vez mais tecnologias estão buscando a melhora na experiência do usuário com vídeos. Existem abordagens imersivas como realidade virtual (RV) e realidade aumentada (RA) que tem como principal objetivo trazer maior imersão dos usuários em uma experiência que une real e virtual. Além disso, existem os vídeos omnidirecionais, esféricos, ou 360°, que permitem que o espectador visualize a cena de dentro dela, podendo olhar ao redor livremente. Ademais, existe a tecnologia tridimensional (3D) com múltiplas vistas, que possibilita uma experiência visual distinta com sensação de profundidade de uma cena. Essa tecnologia atraiu diversos investimentos nos últimos anos, porém, não obteve sucesso e popularização comercial devido a sua falta de naturalidade nas representações visuais, além de possuir certos efeitos colaterais indesejados, como o *Vergence-Accommodation Conflict* (VAC). O VAC ocorre devido à incompatibilidade entre a distância de um objeto virtual 3D e a distância real necessária para o olho humano focar no objeto, podendo causar fadiga ocular, enjôo de movimento e até problemas de visão a longo prazo. Como alternativa a isso, novas tecnologias nesta área estão sendo propostas e a principal são os *light fields*.

Os *light fields* são capazes de representar imagens e suas informações angulares, ou seja, pontos de vistas diferentes de uma mesma cena, trazendo uma representação mais fiel da cena. As principais vantagens dos *light fields* são o processamento pós captura, como mudança de foco e de ponto de vista, extração de dados tridimensionais, mudanças no tamanho do foco, entre outros (IHRKE, 2016). Esta tecnologia é baseada na função plenóptica, que tenta descrever todo o espaço observável em qualquer ponto do espaço e do tempo. O conceito da função plenóptica possibilita um enorme poder de representação, entretanto, na prática, sua representação é reduzida a uma função quadridimensional (4D) $P(r, s, x, y)$, onde (x, y) representam o conjunto de pontos no espaço de uma cena, enquanto que (r, s) representam o conjunto de informações angulares de incidência de luz, isto é, a intensidade de luz incidente em um determinado ponto entre diversos ângulos, como pode ser observado na Figura 1 (a) e 1 (b).

Porém, devido a elevada quantidade de informações visuais que são utilizadas para a representação de *light fields*, são gerados enormes volumes de

dados, trazendo certos desafios em seu armazenamento, transmissão e processamento. Sendo assim, há necessidade de compressão de *light fields*.



(a) Esquema de um *light field*



(b) *Light field* e micro-imagem

Figura 1 – Representações de um *light field*

A partir disso, o grupo JPEG está desenvolvendo soluções para a compressão de *light fields*. O principal objetivo do grupo é prover um formato de compressão eficiente, garantindo alta qualidade de representação do conteúdo, levando em conta fatores como consumo de energia, complexidade computacional, taxas de dados, entre outros. Deste modo, o codificador JPEG Pleno *Verification Model* (VM), o software de referência para a codificação de *light fields* do JPEG, foi desenvolvido para realizar a compressão de *light fields*. Ele traz duas versões, a versão VM 1.0, *Warping and Sparse Prediction* (WaSP) (ASTOLA, 2019), e a versão VM 2.1, *Multidimensional Light-field Encoder using 4D Transforms and Hexadeca-Trees* (MuLE) (CARVALHO, 2018). Este estudo foca na avaliação da eficiência de codificação do software de referência MuLE, que tem como ideia base, a manipulação de *light fields* em quatro dimensões.

2. METODOLOGIA

O grupo JPEG desenvolveu um documento de especificações para padronização dos testes para a codificação de *light fields*, o JPEG Pleno *Light Field Coding Common Test Conditions* (CTCs) (JPEG, 2019). Este documento traz informações de quais os *light fields* devem ser usados para testar soluções, suas especificações, formas de uso e onde obtê-los. Os *light fields* suportados no codificador são os do tipo *High Density Camera Array* (HDCA), onde as imagens são capturadas por um array de câmeras, e os obtidos pela *Lenslet Lytro Illum Camera* (LL), uma câmera com diversas micro-lentes capaz de captar as informações angulares.

Para a codificação, o documento especifica quais são as taxas de compressão alvo a serem obtidas pelo codificador. A taxa de compressão é medida basicamente pelo valor de *bits* por *pixel* (bpp), que é calculado pelo tamanho do arquivo após compressão em *bits*, dividido pela quantidade de *pixels* do mesmo. Quanto menor um valor de bpp para um mesmo *light field* após a codificação, maior a taxa de compressão alcançada. No MuLE, esta taxa de compressão no codificador é controlada pelo parâmetro *lambda*. Este parâmetro é passado para o codificador no momento da execução e pode ser modificado a cada codificação. Entretanto, não é especificado quais valores do parâmetro *lambda* atingem determinadas taxas de compressão para cada *light field*. Deste modo, este trabalho busca determinar tais valores do parâmetro *lambda* para atingir as taxas de compressão alvo propostas e definidas pelo JPEG para o codificador MuLE e avaliar sua eficiência de compressão, ou seja, taxa de compressão e qualidade visual dos *light fields*.

Para atingir estes objetivos, foi necessário realizar a codificação de *light fields* presentes no documento de CTCs com diversos valores diferentes do parâmetro *lambda*. Com o objetivo de automatizar a realização dos testes, foram programados

scripts na linguagem de programação *Python*. A partir disso, foram testados valores para o parâmetro *lambda*, utilizando uma métrica de aceitação de 1% de margem, onde valores até 1% acima ou abaixo do valor de bpp alvo foram considerados. Deste modo, as codificações foram realizadas no *software* de referência MuLE, e a partir de *scripts* para a automatização dos cálculos de bpps os valores dos parâmetros *lambda* foram definidos para os *light fields*.

Após as definições dos *lambda*s, foram realizadas análises de qualidade objetiva dos *light fields* codificados. Esta análise baseia-se na comparação entre o dado original e o dado processado, buscando características que demonstrem os efeitos causados pelo processamento, do mesmo modo em que o sistema visual humano os percebe. Para as análises, foram utilizadas as duas métricas de avaliação objetiva propostas nas CTCs: a *Peak Signal-to-Noise Ratio* (PSNR), que avalia a relação sinal-ruído de pico, e a *Structural Similarity* (SSIM), que mede a similaridade estrutural.

As análises foram realizadas através de *scripts* no *software Matlab*, que facilita operações matriciais. Para cada *light field* codificado, foi gerado seu PSNR e SSIM. O PSNR de um *light field* corresponde à média dos valores de PSNR de todas as vistas individuais ou *sub-aperture images* (SAIs), assim como o SSIM de um *light field* corresponde à média dos valores de SSIM de todas as SAIs. Cada SAI dos *light fields* codificados foi comparada à SAI correspondente do *light field* original, sem codificação, deste modo, a média de todas as SAIs foi calculada e as análises de qualidade objetiva foram geradas e são apresentadas na próxima sessão.

3. RESULTADOS E DISCUSSÃO

Com base na metodologia apresentada, foi possível definir os valores do parâmetro *lambda* para alcançar os bpps alvos definidos pelo JPEG, que são: 0,001, 0,005, 0,02, 0,1 e 0,75, além de determinar como o *lambda* impacta na taxa de compressão. A Figura 2 (a) apresenta a curva *lambda* por bpp em escala logarítmica para o *light field* *Bikes*. Os *light fields* *Danger de Mort*, *Stone Pillars Outside* e *Fountain & Vincent 2* tem comportamento e resultados semelhantes ao *Bikes*, deste modo, não estão apresentados nos gráficos.

A partir desses dados é possível perceber que com o aumento do parâmetro *lambda* a taxa de compressão tende a aumentar, porém, o aumento mais significativo ocorre em *lambda*s menores, até chegar em um bpp próximo de 0,1. A partir desse ponto a taxa de compressão não tende a aumentar drasticamente. Isso pode ser observado no bpp 0,75 – que necessita de um valor de *lambda* de 322 – em relação ao bpp 0,1 – que necessita de um valor de *lambda* de 5800. Nestes dois casos, é possível obter uma grande diferença na taxa de compressão com um aumento pouco significativo do parâmetro *lambda*. Já para o bpp 0,005 – que necessita de um valor de *lambda* de 1250000 – em relação ao bpp 0,001 – que necessita de um valor de *lambda* de 12275000 –, é necessário um *lambda* muito maior para causar uma pequena alteração na taxa de compressão.

Na Figura 2 (b) e 2 (c) é possível observar as métricas de qualidade PSNR e SSIM, respectivamente, para o *light field* *Bikes*. Assim como na curva *lambda* por bpp, os outros *light fields* obtiveram resultados muito semelhantes ao *Bikes* nas avaliações de qualidade objetiva, deste modo, não estão apresentados nos gráficos. A partir destes dados, é possível perceber que com o aumento do valor de bpp, os valores de PSNR e SSIM tendem a aumentar similarmente. Isso se deve ao fato de que maiores taxas de compressão tendem a piorar a qualidade objetiva dos *light fields* codificados. Também é possível analisar que pequenas variações de bpp na região dos bpps menores tendem a impactar mais significativamente na

qualidade do que as mesmas variações na região dos bpps maiores, onde são necessárias grandes variações de bpp para alterar a qualidade significativamente.

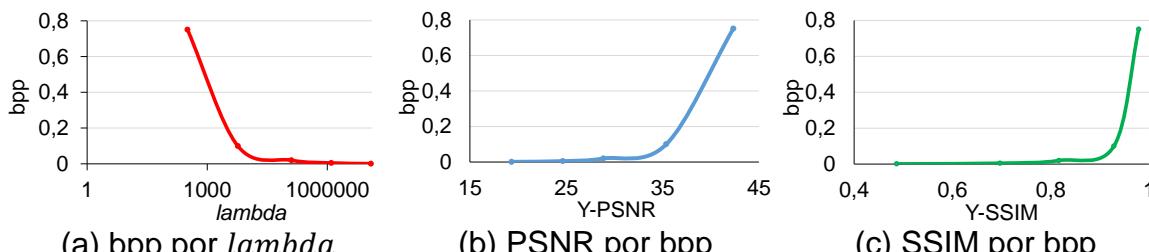


Figura 2 – Resultados de compressão e avaliação de qualidade objetiva

Portanto, levando em conta estes dados, pode-se perceber uma estrita relação entre qualidade e compressão. Com elevadas taxas de compressão pode-se perder na qualidade final dos *light fields*, porém, ao levarmos em conta quantidade de dados, seu armazenamento e transmissão podem ser altamente beneficiados. Levando em conta cenários distintos desejados, pode haver vantagem em maiores taxas de compressão ou maior qualidade de *light fields*.

4. CONCLUSÕES

Este trabalho apresentou a tecnologia de *light fields* e teve como principal objetivo determinar os valores do parâmetro *lambda* para atingir as taxas de compressão alvo propostas e definidas pelo JPEG para o codificador MuLE, além de avaliar sua eficiência de codificação, ou seja, taxa de compressão e qualidade visual dos *light fields*. Como resultados, foram apresentados os valores do parâmetro *lambda* que atingem os bpps alvos para os *light fields* de referência no codificador MuLE, além da sua relação com as taxas de compressão, ou seja, quanto maior os valores do parâmetro *lambda*, menor os valores de *bits per pixel* dos *light fields* codificados. Além disso, foi realizada a análise de qualidade objetiva dos *light fields* codificados com o MuLE com diferentes bpps alvos, gerando valores de PSNR e SSIM, demonstrando uma relação entre qualidade e compressão, ou seja, quanto maior a compressão, menor a qualidade objetiva do *light field*. Consequentemente, trabalhos futuros podem utilizar os resultados apresentados como referência para sua avaliação, utilizando os dados como base e comparação nas codificações de *light fields*. Como trabalhos futuros, pretende-se adicionar previsões no codificador MuLE visando aumento na taxa de compressão.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- CISCO. Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper.** CISCO, 27 fev. 2019. Acessado em 28 ago. 2019. Online. Disponível em: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>
- IHRKE, I.; RESTREPO, J.; MIGNARD-DEBISE, L. Principles of Light Field Imaging. **IEEE Signal Processing Magazine**, 59–69, 2016.
- M. B. de Carvalho *et al.*, A 4D DCT-Based Lenslet Light Field Codec, **IEEE International Conference on Image Processing**, Athens, 2018, pp. 435-439.
- Astola, P., & Tabus, I. (2019). WaSP: Hierarchical Warping, Merging, and Sparse Prediction for Light Field Image Compression. **Proceedings - European Workshop on Visual Information Processing, EUVIP**, 2018-November, 1–6.
- JPEG. **JPEG Pleno - Light Field Coding Common Test Conditions**, Geneva, 2019.